Available online at www.sciencedirect.com
ScienceDirect
Procedia Computer Science 31 (2014) 758 – 765
Procedia Computer Science

2nd International Conference on Information Technology and Quantitative Management, ITQM 2014

Negative attributes and implications in Formal Concept Analysis

J.M. Rodriguez-Jimenez[a], P. Cordero[a,*], M. Enciso[a], A. Mora[a]

[a] *Universidad de Malaga, Campus de Teatinos, Malaga 29017, Spain*

Abstract

The mining of negative attributes from datasets has been studied in the last decade to obtain additional and useful information. There exists an exhaustive study around the notion of negative association rules between sets of attributes. However, in Formal Concept Analysis, the needed theory for the management of negative attributes is in an incipient stage. In this work we present an algorithm, based on the NextClosure algorithm, that allows to obtain mixed implications. The proposed algorithm returns a feasible and complete basis of mixed implications by performing a reduced number of requests to the formal context.

Keywords: Formal concept analysis, attribute implications, data mining, negative attributes.

1. Introduction

Knowledge discovering is nowadays a well established discipline focussed on the development of tools and techniques to reveal useful information hidden in big amount of data. Its main goal is to detect patterns to improve decision making and is approached using pattern recognition, clustering, association and classification. Part of these patterns are expressed as implications (or association rules) which allow us to address information using a formal notation and to manage them syntactically by using logic.

Thus, implications are formulas in the form $A \to B$ where A and B are subsets of a certain set (universe) of attributes M. Both subsets of attributes are considered to be conjunctive cubes, i.e. $A = a_1 \wedge \ldots \wedge a_n$ and $B = b_1 \wedge \ldots \wedge b_m$. As we mention above, its main advantage is the existence of logics developed to specify and manage sets of implications. The pioneer of these logics was the one introduced by W.W.Armstrong[1], which was proven to be sound and complete.

In this work we use implications in the area of formal concept analysis[2] and assume the common interpretation in this environment: given a formal context $\mathbb{K}$ over a set of attributes Ω, the implication $A \to B$ asserts that any object which have all the A attributes, also has all the B attributes. Although we focus on data mining techniques to extract implications in formal concept analysis, this problem is similar to the extraction of functional dependencies or association rules from an arbitrary data set.

* P. Cordero. Tel.: +34-952-137165 ; fax: +34-952-13-2766.
E-mail address: pcordero@uma.es

1877-0509
doi:10.1016/j.procs.2014.05.325

One of the former researchers who points out the importance of this problem in datasets was H. Mannila[3] and it was also studied by other researches from the database areas like S. Navathe[4]. In this work we will address the mining of implications with negation. The extended implications allow us to relate items which conflict with each other. While classical implications express that *"cyclist with short and sharp accelerations are great sprinters"*, implications with negations allow us to express that *"cyclist with short and sharp accelerations are not great climbers"*.

Since implication formulas are built using a binary connective which relates two conjunctive clauses, negation is only considered at the attribute level, i.e. the negation $\overline{A}$ is considered the conjunctive clause of its negated attributes: $\overline{a_1} \wedge \ldots \wedge \overline{a_n}$. Notice that extended implications cannot be considered as the negation of a classical implication. We do not want to express that a certain implication does not holds but the evidence of the absence of a certain attribute.

Mining implication with negation is a well known problem which has been exhaustively studied and some founding papers may be cited[5,6]. In this paper we are going to focus on this problem in the framework of formal concept analysis. In this area, some authors propose the mining of these implications from the apposition of the context and its negation $(\mathbb{K}|\overline{\mathbb{K}})$[7]. Although it allows to use previously developed methods, as R. Missaoui et.al. shown in[8] real applications use to have sparse data in the context $\mathbb{K}$ whereas dense data in $\overline{\mathbb{K}}$ (or viceversa), and therefore *"generate a huge set of candidate itemsets and a tremendous set of uninteresting rules"*. This situation has been also mentioned in[9], as an argument to work in the development of more efficient methods, and in[7], where the authors propose to extract a basis of generalized rules and, from this basis, the whole set of generalized rules may be extracted.

In our opinion the works of R. Missaoui et al. constitutes a solid approach to this problem[10,11]. They mining both positive and negative implications (exact association rules) from a (positive) attributes formal context. In these works they deal in two separate steps with positive and negative implications to generate in a further step a set of mixed implications. The problem was addressed by using a set of inference rules which allows to produce a basis with useful and feasible set of implications. Recently, in[8], they propose the generation of the set of mixed implications in a direct way by introducing the concept of *key*, i.e. a subset of attributes which constitutes a minimal generator of the built formal concept in $(\mathbb{K}|\overline{\mathbb{K}})$, which corresponds exactly with the classical notion of key of databases.

In this work, we develop a method to mine mixed implications from a formal context but instead of using the large $(\mathbb{K}|\overline{\mathbb{K}})$ context, we preserve the original one $(\mathbb{K})$ and provide an extension of the closure operators. The method is strongly based on a set of inference rules for generalized implications and it generates a basis of implications from which all generalized implications may be extracted. Later, in Section 2 we will introduce some notation and previous results to introduce our method in Section 3. The application of the method traverses the set of itemsets following an extended lectic order that we will introduce in the paper. The paper also includes two illustrative examples in Section 4 which shows its benefits. Since we consider this work as a first step of a promising line, we end with a conclusion and further works to indicate a medium-term guideline in Section 5.

2. Preliminaries

In this section, the basic notions related with Formal Concept Analysis (FCA)[12] and attribute implications are briefly presented. See[13] for a more detailed explanation. A **formal context** is a triple $\mathbb{K} = \langle G, M, I\rangle$ where G and M are finite non-empty sets and $I \subseteq G \times M$ is a binary relation. The elements in G are named objects, the elements in M attributes and $\langle g, m\rangle \in I$ means that the object g has the attribute m. From this triple, two mappings $\uparrow$: $2^G \to 2^M$ and $\downarrow$: $2^M \to 2^G$, named concept-forming operators, are defined as follows: for any $X \subseteq G$ and $Y \subseteq M$,

$$X^{\uparrow} = \{m \in M | \text{ for each } g \in X : \langle g, m\rangle \in I\} \tag{1}$$

$$Y^{\downarrow} = \{g \in G \mid \text{ for each } m \in Y : \langle g, m\rangle \in I\} \tag{2}$$

$X^{\uparrow}$ is the subset of all attributes shared by all the objects in X and $Y^{\downarrow}$ is the subset of all objects that have the attributes in Y. The pair $(\uparrow, \downarrow)$ constitutes a Galois connection between 2^G and 2^M and, therefore, both compositions are closure operators.

A pair of subsets $\langle X, Y\rangle$ with $X \subseteq G$ and $Y \subseteq M$ such $X^{\uparrow} = Y$ and $Y^{\downarrow} = X$ is named a **formal concept**. X is named the **extent** and Y the **intent** of the concept. These extents and intents coincide with closed sets wrt the closure operators because $X^{\uparrow\downarrow} = X$ and $Y^{\downarrow\uparrow} = Y$. Thus, the set of all the formal concepts is a lattice, named **concept lattice**, with the relation

$$\langle X_1, Y_1\rangle \leq \langle X_2, Y_2\rangle \text{ if and only if } X_1 \subseteq X_2 \text{ (or equivalently, } Y_2 \subseteq Y_1) \tag{3}$$

The concept lattice can be characterized in terms of attribute implications. An **attribute implication** is an expression $A \to B$ where $A, B \subseteq M$ and it holds in a formal context if $A^{\downarrow} \subseteq B^{\downarrow}$. That is, any object that has all the attributes in A has also all the attributes in B. It is well known that the sets of attribute implications that are satisfied by a context satisfies the Armstrong's Axioms:

`[Ref]` Reflexivity: If $B \subseteq A$ then $\vdash A \to B$.

`[Augm]` Augmentation: $A \to B \vdash A \cup C \to B \cup C$.

`[Trans]` Transitivity: $A \to B, B \to C \vdash A \to C$.

A set of implications $\mathfrak{B}$ is an **implicational base** for $\mathbb{K}$ if: (1) any implication from $\mathfrak{B}$ holds in $\mathbb{K}$ and (2) any implication that $\mathbb{K}$ satisfies follows (can be inferred) by using Armstrong's Axioms from $\mathfrak{B}$.

One of the most cited kind of basis is the so-called Duquenne-Guigues (or stem) base[14]. The premises of the implications in the Duquenne-Guigues base are pseudo-intents: $P \subseteq M$ is a **pseudo-intent** if P is not an intent ($P^{\downarrow\uparrow} \neq P$) and $Q^{\downarrow\uparrow} \subseteq P$ holds for every pseudo-intent $Q \subsetneq P$. The Duquenne-Guigues base for $\mathbb{K}$ is

$$\{P \to (P^{\downarrow\uparrow} \smallsetminus P) \mid P \text{ is a pseudo-intent for } \mathbb{K}\} \tag{4}$$

and satisfies that its cardinality is minimum among all the bases. It is well-known the **NextClosure** Algorithm[13] that computes all the pseudo-intents and intents, and therefore the Duquenne-Guigues base for a context. This algorithm is based in the **lectic order** among sets of attributes that coincides with the usual order for binary numbers when set of attributes are represented by bit-maps.

3. Computing implications with negative and positive attributes

As mentioned in the introduction, the aim of this work is to give an algorithm for computing generalized attribute implications (implications in which attributes can appear asserted or negated) directly from a formal context. One trivial approach consists on duplicating the formal context with the opposite context. That is, given a formal context $\mathbb{K} = \langle G, M, I\rangle$, the opposite context is defined as $\overline{\mathbb{K}} = \langle G, \overline{M}, \overline{I}\rangle$ where $\overline{M} = \{\overline{m} \mid m \in M\}$, and $\overline{I} = \{\langle g, \overline{m}\rangle \mid g \in G, m \in M, \langle g, m\rangle \notin I\}$. Thus, $\overline{m}$ is read as "not m" because $\langle g, \overline{m}\rangle \in \overline{I}$ if and only if $\langle g, m\rangle \notin I$, i.e. the object g has not the attribute m.

The attributes in M are said to be positive whereas the elements of $\overline{M}$ are named negative attributes. Obviously, implications among positive attributes can be obtained from $\mathbb{K}$ and implications among sets of negative attributes are obtained from $\overline{\mathbb{K}}$. However, it is not possible to obtain implications related to positive and negative attributes simultaneously. A coarse solution can be obtained by juxtaposition of both contexts, the first one and its opposite: $(\mathbb{K}|\overline{\mathbb{K}}) = \langle G, M \cup \overline{M}, I \cup \overline{I}\rangle$.

Example 1. *Let us consider the context $\mathbb{K} = \langle G, M, I\rangle$ (see[10]), in which the objects set is $G = \{x_1, x_2, x_3, x_4\}$, the attributes set is $M = \{a, b, c, d\}$ and whose binary relation I is depicted in Table 1. The opposite formal context $\overline{\mathbb{K}} = \langle G, \overline{M}, \overline{I}\rangle$ is depicted in Table 2. Finally, the composed formal context $(\mathbb{K}|\overline{\mathbb{K}}) = \langle G, M \cup \overline{M}, I \cup \overline{I}\rangle$ is depicted in Table 3.*

	a	b	c	d
x_1	×			×
x_2		×		×
x_3	×	×		
x_4		×	×	×

Table 1. Formal context $\mathbb{K} = \langle G, M, I\rangle$.

Obviously, any attribute implication which is valid in $\mathbb{K}$ is valid in $(\mathbb{K}|\overline{\mathbb{K}})$ also. For example, the implication $bc \to d$ holds in the context $\mathbb{K}$ and, therefore, in $(\mathbb{K}|\overline{\mathbb{K}})$. On the other hand, any valid implication in $\overline{\mathbb{K}}$ is valid in $(\mathbb{K}|\overline{\mathbb{K}})$ also. See, for example, $\overline{b} \to \overline{c}$. However, there exist implications that hold in $(\mathbb{K}|\overline{\mathbb{K}})$ and can not be obtained from $\mathbb{K}$ or from $\overline{\mathbb{K}}$. It occurs, for example, for $a \to \overline{c}$. They are mixed implications in which negative and positive attributes take part.

	$\overline{a}$	$\overline{b}$	$\overline{c}$	$\overline{d}$
x_1		×	×	
x_2	×		×	
x_3			×	×
x_4	×			

Table 2. Formal context $\overline{\mathbb{K}} = \langle G, \overline{M}, \overline{I}\rangle$ which is the opposite of $\mathbb{K}$.

	a	b	c	d	$\overline{a}$	$\overline{b}$	$\overline{c}$	$\overline{d}$
x_1	×			×		×	×	
x_2		×		×	×		×	
x_3	×	×					×	×
x_4		×	×	×	×			

Table 3. Formal context $(\mathbb{K}|\overline{\mathbb{K}}) = \langle G, M \cup \overline{M}, I \cup \overline{I}\rangle$ which is the juxtaposition of $\mathbb{K}$ and its opposite $\overline{\mathbb{K}}$.

A rude solution is to compute all attribute implications from $(\mathbb{K}|\overline{\mathbb{K}})$. However, this solution does not take advantage of the existent relation between attributes and its negations. Moreover, the number of attribute sets to be explored increases from $2^{|M|}$ to $4^{|M|}$. As far as we know, all related works go in the line to reduce the cost of computing a base of implications from $(\mathbb{K}|\overline{\mathbb{K}})$. However, with our approach, we try to obtain a base directly from $\mathbb{K}$ by extending classical results and algorithms. First, we extend the definitions of concept-forming operators, formal concepts and attribute implications.

Definition 1. *Let $\mathbb{K} = \langle G, M, I\rangle$ be a formal context. We define the operators $\Uparrow: 2^G \to 2^{M\cup\overline{M}}$ and $\Downarrow: 2^{M\cup\overline{M}} \to 2^G$ as follows: for each $X \subseteq G$ and each $Y \subseteq M \cup \overline{M}$,*

$$X^{\Uparrow} = \{m \in M |\, \text{for all } g \in X : \langle g, m\rangle \in I\} \cup \{\overline{m} \in \overline{M} |\, \text{for all } g \in X : \langle g, m\rangle \notin I\} \tag{5}$$

$$Y^{\Downarrow} = \{g \in G |\, \text{for all } m \in Y : \langle g, m\rangle \in I\} \cap \{g \in G |\, \text{for all } \overline{m} \in Y : \langle g, m\rangle \notin I\} \tag{6}$$

*which constitute a Galois connection between $(2^G, \subseteq)$ and $(2^{M\cup\overline{M}}, \subseteq)$. A **mixed formal concept** is a pair of subsets $\langle X, Y\rangle$ with $X \subseteq G$ and $Y \subseteq M \cup \overline{M}$ such $X^{\Uparrow} = Y$ and $Y^{\Downarrow} = X$.*

*Given two subsets $A, B \subseteq M \cup \overline{M}$, the context $\mathbb{K}$ satisfies a **mixed attribute implication** $A \to B$, denoted by $\mathbb{K} \models A \to B$, if $A^{\Downarrow} \subseteq B^{\Downarrow}$.*

At this point, some questions about the notation need to be fixed. From now on, the set of all the attributes is denoted by M, and its elements by the letter m possibly with subindexes. The elements in $M \cup \overline{M}$ are going to be denoted by the first letters in the alphabet: a, b, c... So, the symbols a, b, c... could represent positive or negative attributes. Capital letters A, B, C ... denote subsets of $M \cup \overline{M}$. If $A \subseteq M \cup \overline{M}$, then $\overline{A}$ denotes the set of the opposite of attributes in A. That is, $\overline{A} = \{\overline{a} \mid a \in A\}$ where $\overline{\overline{a}} = a$. Moreover, for $A \subseteq M \cup \overline{M}$, the following sets are defined:

$$\text{Pos}(A) = \{m \in M \mid m \in A\};\ \text{Neg}(A) = \{m \in M \mid \overline{m} \in A\};\ \text{Tot}(A) = \text{Pos}(A) \cup \text{Neg}(A) \tag{7}$$

and, therefore, $\text{Pos}(A), \text{Neg}(A), \text{Tot}(A) \subseteq M$.

Obviously, Armstrong's Axioms can be used for reasoning with mixed implications because they are fulfilled by the formal context $(\mathbb{K}|\overline{\mathbb{K}})$. However, we need to extend the axiomatic system to capture the specific behavior of each attribute and its negation. Thus, in this section we define new axioms and inference rules in which attributes and its negations appear.

Theorem 1. *Let $\mathbb{K} = \langle G, M, I\rangle$ be a formal context. For all $a, b \in M \cup \overline{M}$ and $A \subseteq M \cup \overline{M}$, the following assertions hold:*

1. $\mathbb{K} \models a\overline{a} \to M\overline{M}$
2. *If* $\mathbb{K} \models Aa \to b$ *then* $\mathbb{K} \models A\overline{b} \to \overline{a}$

Proof. The first assertion is trivial because, in $\mathbb{K}$, we have that $(a\overline{a})^{\Downarrow} = \varnothing = (M\overline{M})^{\Downarrow}$.

For the second one, assume $(Aa)^{\Downarrow} \subseteq b^{\Downarrow}$. Since $(A\overline{b})^{\Downarrow} = A^{\Downarrow} \cap \overline{b}^{\Downarrow} \subseteq \overline{b}^{\Downarrow}$ and $A^{\Downarrow} \cap a^{\Downarrow} = (Aa)^{\Downarrow} \subseteq b^{\Downarrow}$, we have that $A^{\Downarrow} \cap a^{\Downarrow} \cap \overline{b}^{\Downarrow} \subseteq b^{\Downarrow} \cap \overline{b}^{\Downarrow} = \varnothing$. Therefore, $(A\overline{b})^{\Downarrow} = A^{\Downarrow} \cap \overline{b}^{\Downarrow} \subseteq G \smallsetminus a^{\Downarrow} = \overline{a}^{\Downarrow}$. □

Therefore, the following axioms are aded to the Armstrong's axioms: for all $a, b \in M \cup \overline{M}$ and $A \subseteq M \cup \overline{M}$,

`[Cont]` Contradiction: $\vdash a\overline{a} \to M\overline{M}$.

`[Rft]` Reflection: $Aa \to b \vdash A\overline{b} \to \overline{a}$.

The closure of an attribute set A wrt a set of mixed attribute implications $\mathfrak{B}$, denoted as A^{+}, is defined as the biggest set such that $A \to A^{+}$ can be inferred from $\mathfrak{B}$ by using Armstrong's Axioms plus `[Cont]` and `[Rft]`. Therefore, a mixed implication $A \to B$ can be inferred from $\mathfrak{B}$ if and only if B is a subset of the closure of A, i.e. $B \subseteq A^{+}$.

Moreover, a Closure algorithm for mixed implications can be defined as follows:

Algorithm 1: Closure algorithm for mixed attribute implications

Data: $\mathfrak{B}$ being a set of mixed implications, A being a set of (positive or negative) attributes
Result: A^{+} (the closure of A)

begin
  repeat
    $A_{\texttt{old}} := A$;
    if $A \cap \overline{A} \neq \varnothing$ **then**
      $A := M\overline{M}$
    else
      foreach $B \to C \in \mathfrak{B}$ **do**
        if $B \subseteq A$ **then**
          $A := A \cup C$
        if $B \smallsetminus A$ *is a singleton, i.e.* $B \smallsetminus A = \{a\}$, *and* $A \cap \overline{C} \neq \varnothing$ **then**
          $A := A \cup \{\overline{a}\}$
  until $A = A_{\texttt{old}}$ *or* $A = M\overline{M}$;
  return A
end

Note that, for a set of attributes $A \subseteq M \cup \overline{M}$, if $A \cap \overline{A} \neq \varnothing$ or, equivalently, $\mathrm{Pos}(A) \cap \mathrm{Neg}(A) \neq \varnothing$, then $A^{\Downarrow} = \varnothing$, $A^{\Downarrow\Uparrow} = M \cup \overline{M}$ and $A^{+} = M\overline{M}$. Therefore, any implication with A in the left hand side can be inferred from an axiom (see `[Cont]`). Thus, if this implication appears in a base, another equivalent base with less cardinality can be obtained by removing this implication. The algorithm that we propose in this work, following the scheme of NextColure algorithm, traverses the set of subsets of $M \cup \overline{M}$ that can be a left hand side of a non-trivial implication. So, the algorithm considers only sets $A \subseteq M \cup \overline{M}$ such that $\mathrm{Pos}(A) \cap \mathrm{Neg}(A) = \varnothing$ that will be named **consistent** sets. The set of consistent sets are going to be denoted by $\mathbb{C}$tts.

$$\mathbb{C}\mathrm{tts} = \{A \subseteq M \cup \overline{M} \mid \mathrm{Pos}(A) \cap \mathrm{Neg}(A) = \varnothing\} \tag{8}$$

Therefore, if $A \in \mathbb{C}$tts then $|A| \leq |M|$ and, in the case of $|A| = |M|$, we have that $\mathrm{Tot}(A) = M$. Sets $A \in \mathbb{C}$tts such that $\mathrm{Tot}(A) = M$ will be named **full** sets.

Lemma 1. *Given $A \in \mathbb{C}$tts and an implication base $\mathfrak{B}$, if A is full then A is closed or $A^{+} = M\overline{M}$.*

The following theorem is a consequence of Theorem 1, the axiomatic system and, particularly, from Algorithm 1.

Theorem 2 (Characterization of closed sets)**.** *Let $\mathfrak{B}$ be a base of mixed implications for a context $\mathbb{K}$. A set $A \in \mathbb{C}$tts is closed wrt $\mathfrak{B}$ (i.e. $A^{+} = A$) if and only if the following conditions hold:*

1. *For all $B \to C \in \mathfrak{B}$, if $B \subseteq A$ then $C \subseteq A$.*

2. *For all $B \to C \in \mathfrak{B}$, if $B \smallsetminus A$ is the singleton $\{a\}$ and $A \cap \overline{C} \neq \varnothing$ then $\overline{a} \in A$.*

Function `Closed(`A,$\mathfrak{B}$`)`: boolean

Data: $A \in \mathbb{C}$tts, and $\mathfrak{B}$ being a set of mixed implications.
Result: '`true`' if A is closed wrt $\mathfrak{B}$ or '`false`' otherwise.

begin
    foreach $B \to C \in \mathfrak{B}$ **do**
        if $B \subseteq A$ *and* $C \nsubseteq A$ **then**
            exit and **return** `false`
        if $B \smallsetminus A = \{a\}$, $A \cap \overline{C} \neq \varnothing$, *and* $\overline{a} \notin A$ **then**
            exit and **return** `false`
    return `true`
end

The last step to introduce the algorithm for computing a mixed implication base from a formal context is the extension to $2^{M \cup \overline{M}}$ of the classical lectic order $<$ defined in 2^M. To do this, we are going to give a characterization of sets in $\mathbb{C}$tts in terms of subsets of M. There is a one-to-one relation between sets in $\mathbb{C}$tts and pairs of sets $\langle X, Y\rangle$ such that $X \subseteq Y \subseteq M$. If $A \in \mathbb{C}$tts then $\mathrm{Neg}(A) \subseteq \mathrm{Tot}(A) \subseteq M$. On the other hand, if $X \subseteq Y \subseteq M$ then $(Y \smallsetminus X) \cup \overline{X} \in \mathbb{C}$tts. Moreover,

$$A = \Big(\mathrm{Tot}(A) \smallsetminus \mathrm{Neg}(A)\Big) \cup \overline{\mathrm{Neg}(A)} \tag{9}$$

This one-to-one relation allows to extend the lectic order to $\mathbb{C}$tts as follows: for all $A, B \in \mathbb{C}$tts, the set A is previous to B, denoted as $A \ll B$, if one of the following conditions holds:

1. $\mathrm{Tot}(A) < \mathrm{Tot}(B)$
2. $\mathrm{Tot}(A) = \mathrm{Tot}(B)$ and $\mathrm{Neg}(A) < \mathrm{Neg}(B)$

Observe that, if A is a non-full set and B is a full set then $A \ll B$. On the other hand, any implication with a full set in its left hand side is an axiom (see Lemma 1) and, therefore, full sets are dispensable in the exploration. In the algorithm we will traverse the set $\mathbb{C}$tts in the $\ll$ order until the first full set appears.

Algorithm 3: Mixed Implications Mining

Data: $\mathbb{K} = \langle G, M, I\rangle$
Result: $\mathfrak{B}$ basis of implications

begin
    $\mathfrak{B} := \varnothing$;
    $Y := \varnothing$;
    while $Y < M$ **do**
        foreach $X \subseteq Y$ **do**
            $A := (Y \smallsetminus X) \cup \overline{X}$;
            if `Closed(`A, $\mathfrak{B}$`)` **then**
                $C := A^{\Downarrow\Uparrow}$;
                if $A \neq C$ **then** $\mathfrak{B} := \mathfrak{B} \cup \{A \to C \smallsetminus A\}$
        $Y :=$ `Next(`Y`) // i.e. the successor of` Y `in the classical lectic order on` 2^M
    return $\mathfrak{B}$
end

---

4. Illustrative examples

In this section we show how our method works in two different scenarios. In the first example we compare the method with a data mining method which does not infer negative information while in the second it is contrasted with a similar method which extract mixed implications.

Example 2. *In this first example we compute a base of mixed implications for one of the most cited example in the literature of formal concept analysis, presented in*[13]*. The context represents developing countries as objects and the attributes are supranational associations of countries, so that the context depicts the belonging of countries to these institutions. It is built with 130 countries and 6 attributes: Group of 77, Non-alligned, LDC (Least Developed Countries), MSAC (Most Seriously Affected Countries), OPEC (Organization of Petrol Exporting Countries) and ACP (African, Caribean and Pacific Countries). The implications which makes up the Duquenne-Guigues base are the following:*

$$\begin{array}{l} \mathit{OPEC} \;\rightarrow\; \mathit{Group\ of\ 77,\ Non\text{-}alligned} \\ \mathit{MSAC} \;\rightarrow\; \mathit{Group\ of\ 77} \\ \mathit{Non\text{-}alligned} \;\rightarrow\; \mathit{Group\ of\ 77} \\ \mathit{Group\ of\ 77,\ Non\text{-}alligned,\ MSAC,\ OPEC} \;\rightarrow\; \mathit{LLDC,\ ACP} \\ \mathit{Group\ of\ 77,\ Non\text{-}alligned,\ LLDC,\ OPEC} \;\rightarrow\; \mathit{MSAC,\ ACP} \end{array} \tag{10}$$

The basis generated by our method is built with the following mixed attribute implications:

$$\begin{array}{l} \mathit{OPEC} \rightarrow \mathit{Group\ of\ 77,\ Non\text{-}alligned},\ \overline{\mathit{LLDC}},\ \overline{\mathit{MSAC}} \\ \mathit{MSAC},\ \overline{\mathit{OPEC}} \rightarrow \mathit{Group\ of\ 77} \\ \overline{\mathit{LLDC}},\ \mathit{ACP} \rightarrow \mathit{Group\ of\ 77} \\ \overline{\mathit{Non\text{-}alligned}},\ \mathit{LLDC},\ \overline{\mathit{MSAC}},\ \overline{\mathit{OPEC}} \rightarrow \overline{\mathit{Group\ of\ 77}} \\ \mathit{Non\text{-}alligned} \rightarrow \mathit{Group\ of\ 77} \\ \mathit{Group\ of\ 77},\ \overline{\mathit{LLDC}},\ \mathit{MSAC},\ \overline{\mathit{OPEC}},\ \mathit{ACP} \rightarrow \mathit{Non\text{-}alligned} \end{array} \tag{11}$$

As a conclusion, in the execution of the method the formal context is checked only 124 times whereas the cardinality of $\mathbb{C}$tts is 729. The number of implications in the mixed implication basis is 6, just one implication more than the classical Duquenne-Guigues basis. In our opinion although mixed implications held in a context are significantly greater than classical implications, our method does not infer a lot of mixed implications which may be useless but it produces a feasible and useful information.

Example 3. *In this second example, we compare the basis built with our method with that obtained with the algorithm introduced in*[10]*. For the formal context given by Table 1, the Missaoui et al. algorithm renders the following mixed implication basis:*

$$\mathfrak{B}_1 = \{\overline{d} \rightarrow ab\overline{c}, c \rightarrow \overline{a}bd, \overline{b} \rightarrow a\overline{c}d, bd \rightarrow \overline{a}, \overline{a} \rightarrow bd, a \rightarrow \overline{c}, ad \rightarrow \overline{b}\overline{c}, ab \rightarrow \overline{c}\overline{d}, \overline{ac} \rightarrow bd, b\overline{c}d \rightarrow \overline{a}\} \tag{12}$$

However, the algorithm that we propose in this paper renders the following basis:

$$\mathfrak{B}_2 = \{\overline{d} \rightarrow ab\overline{c}, cd \rightarrow \overline{a}b, bd \rightarrow \overline{a}, \overline{b}\overline{c}d \rightarrow a, a \rightarrow \overline{c}, \overline{a}d \rightarrow b\} \tag{13}$$

Our method checks the table 14 *times while the number of possible subsets of attributes is* 81 *and our basis is smaller than the one obtained in the original example: it has* 6 *implications opposite to the* 10 *implications obtained in the original example.*

5. Conclusions

In this work we have introduced a method to mine generalized implications from a formal context. The method allows to generate a basis of implications which characterizes, in the form of implications, the relations among positive and negative attributes, i.e. they connect the presence of attributes in the context and the absence of some counterparts.

The main novelty of our approach is the extension of the closure operators to avoid the growth of the formal context. Thus, we directly work with the original formal context $\mathbb{K}$ and not with the modified $(\mathbb{K}|\overline{\mathbb{K}})$. The introduced method is inspired in the NextClosure method where the lectic order has been extended.

We consider this work as a first step of a wider and deeper line whose main issues are the following:

- A development of a deep algebraic study to built Galois connection considering negative attributes.
- The introduction of a logic for mixed implications following the line introduced in [15] with the development of the Simplification Logic.
- To investigate the properties of the notion of basis for mixed implications generated by the method and make a precise matching of this notion w.r.t. the minimality properties and others, following the study developed in [16] for canonical basis.
- To consider other alternatives to NextClosure to develop new methods, incorporating the extended closure operator and inference rules to different approaches like Titanic [17] or R. Missaoui method [8]. In this line we propose to carry out an efficient implementation of different methods and a precise benchmark to show their real behavior despite of their theoretical (and hard) complexity.

Acknowledgements

Supported by grant TIN2011-28084 of the Science and Innovation Ministry of Spain, co-funded by the European Regional Development Fund (ERDF).

References

1. Armstrong, W.. Dependency structures of data base relationships. *Proc IFIP Congress North Holland, Amsterdam* 1974;:580–583.
2. Ganter, B., Wille, R.. Formal concept analysis. *Mathematical Foundations* 1999;.
3. Mannila, H., Toivonen, H., Verkamo, A.I.. Efficient algorithms for discovering association rules. In: *KDD Workshop*. 1994, p. 181–192.
4. Savasere, A., Omiecinski, E., Navathe, S.B.. An efficient algorithm for mining association rules in large databases. In: *VLDB*. 1995, p. 432–444.
5. Brin, S., Motwani, R., Silverstein, C.. Beyond market baskets: Generalizing association rules to correlation. In: *Proceedings of the 1997 ACM SIGMOD international conference on Management of Data (SIGMOD'97)*. ACM Press; 1997, p. 265–276.
6. Savasere, A., Omiecinski, E., Navathe, S.. Mining for strong negative associations in a large database of customer transactions. In: *ICDE*. IEEE Computer Society; 1998, p. 494–502.
7. Gasmi, G., Yahia, S.B., Nguifo, E.M., Bouker, S.. Extraction of association rules based on literalsets. In: *DaWaK*. 2007, p. 293–302.
8. Missaoui, R., Nourine, L., Renaud, Y.. Computing implications with negation from a formal context. *Fundam Inform* 2012;**115**(4):357–375.
9. Boulicaut, J.F., Bykowski, A., Jeudy, B.. Towards the tractable discovery of association rules with negations. In: *FQAS*. 2000, p. 425–434.
10. Missaoui, R., Nourine, L., Renaud, Y.. Generating positive and negative exact rules using formal concept analysis: Problems and solutions. In: *ICFCA*. 2008, p. 169–181.
11. Missaoui, R., Nourine, L., Renaud, Y.. An inference system for exhaustive generation of mixed and purely negative implications from purely positive ones. In: *CLA*. 2010, p. 271–282.
12. Wille, R.. Restructuring lattice theory: an approach based on hierarchies of concepts. In: *Rival, I. (ed.): Ordered Sets*. Boston; 1982, p. 445–470.
13. Ganter, B.. Two basic algorithms in concept analysis. *Technische Hochschule, Darmstadt* 1984;.
14. Guigues, J., Duquenne, V.. Familles minimales d implications informatives resultant d un tableau de donnees binaires. *Mathematiques et Sciences Sociales* 1986;**95**:5–18.
15. Cordero, P., Enciso, M., Mora, A., Ojeda-Aciego, M.. Computing minimal generators from implications: a logic-guided approach. In: *CLA*. 2012, p. 187–198.
16. Bertet, K., Monjardet, B.. The multiple facets of the canonical direct unit implicational basis. *Theor Comput Sci* 2010;**411**(22-24):2155–2166.
17. Stumme, G., Taouil, R., Bastide, Y., Pasquier, N., Lakhal, L.. Computing iceberg concept lattices with titanic. *Data and Knowledge Engineering* 2002;**42**(2):189–222.